

The Claims

Claims 1-34 (Canceled).

35. (Currently amended) A method as recited in claim 43 ~~[[34]]~~, wherein the periodicity features include a noise frame ratio that identifies a ratio of noise frames to non-noise frames in the plurality of frames.

36. (Original) A method as recited in claim 35, wherein the classifying comprises classifying at least the portion as environment sound if the noise frame ratio exceeds a threshold value.

37. (Currently amended) A method as recited in claim 43 ~~[[34]]~~, further comprising:

extracting, from the plurality of frames, a band periodicity for each of a plurality of bands of the audio signal and a full band periodicity that is a concatenation of the band periodicities for each of the plurality of bands; and

wherein the classifying comprises classifying at least the portion as environment sound if a band periodicity of a first of the plurality of bands is less than the first threshold a band periodicity of a second of the plurality of bands is less than the second threshold.

38. (Currently amended) A method as recited in claim 43 ~~[[34]]~~, wherein the periodicity features include a band periodicity for each of a plurality of bands of the audio signal.

39. (Original) A method as recited in claim 38, further comprising:
extracting a full band periodicity from the plurality of frames that is a concatenation of the band periodicities for each of the plurality of bands; and
wherein the classifying comprises classifying at least the portion as environment sound if the full band periodicity exceeds a threshold value.

40. (Currently amended) A method as recited in claim 43 ~~[[34]]~~, further comprising:
extracting a spectrum flux feature from the plurality of frames; and
wherein the classifying comprises classifying at least the portion as either music or environment sound based at least in part on the periodicity feature and the spectrum flux feature.

41. (Canceled).

42. (Canceled).

43. (Currently amended) A method ~~as recited in claim 42~~, further comprising:
receiving an audio signal;

separating the audio signal into a plurality of portions;

classifying each of the plurality of portions, based at least in part on periodicity features of the portion, as one of: speech, music, silence, and environment sound;

extracting line spectrum pairs from each of the plurality of frames;

generating an input Gaussian Model corresponding to the plurality of frames based on the extracted line spectrum pairs;

identifying one of the plurality of trained Gaussian Models that is closest to the input Gaussian Model;

determining a distance between the input Gaussian Model and the closest trained Gaussian Model;

classifying at least the portion as one of music, silence, or environment sound if the distance is greater than a first threshold value;

determining an energy distribution of the plurality of frames in a first bandwidth; and

classifying at least the portion as one of music, silence, or environment sound if the distance is greater than a second threshold value and the energy distribution of the plurality of frames in the first bandwidth is less than a third threshold value, wherein the second threshold value is less than the first threshold value.

44. (Original) A method as recited in claim 43, further comprising:

determining an energy distribution of the plurality of frames in a second bandwidth; and

classifying at least the portion as one of music, silence, or environment sound if the distance is greater than a fourth threshold value and the energy distribution of the plurality of frames in the second bandwidth is less than a fifth threshold value, wherein the fourth threshold value is less than the first threshold value.

45. (Original) A method as recited in claim 44, further comprising otherwise classifying at least the portion as speech.

Claims 46-50 (Canceled).